Canonical Bases in Linear Programming*

Adi Ben-Israel†

Department of Mathematical Sciences
University of Delaware
Newark, Delaware 19716
and
Institutes of Economics and Informatics
University of Bergen
N-5000 Bergen, Norway

Submitted by Richard A. Brualdi

ABSTRACT

This paper gives an alternative, unified development of the primal and dual simplex methods for maximizing

$$\mathbf{c}^T\mathbf{x} \qquad \text{subject to} \quad A\mathbf{x}=\mathbf{b}, \quad \mathbf{x}\geqslant 0.$$

The calculations are described in terms of certain canonical bases for the null space of A and the range space of A^T. The vectors of these bases are edges of the polyhedron in question at the given basic feasible solution.

0. NOTATION

The common notation of linear algebra is used. In particular:

$R^{m\times n}$ $[R_r^{m\times n}]$ = the $m\times n$ real matrices [of rank r].

*A revision of Report CMI-852555-7, The Chr. Michelsen Institute, N-5036 Fantoft, Norway, Aug. 1985.

†Supported by the National Science Foundation and by the Royal Norwegian Council for Industrial and Scientific Research.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/88/$3.50

For any $A \in R^{m\times n}$:

$$N(A) = \text{the null space of } A,$$
$$R(A^T) = \text{the range of its transpose } A^T.$$

For any two integers j, k:

$$\overline{j,k} := \begin{cases} \{j, j+1, \dots, k\} & \text{if } j \leqslant k, \\ \varnothing & \text{otherwise.} \end{cases}$$

For any $A \in R^{m\times n}$ and subsets $\mathscr{I} \subset \overline{1,m}$, $\mathscr{J} \subset \overline{1,n}$:

$A[\mathscr{I},\ ]$ = the submatrix consisting of rows in $\mathscr{I}$; $A[i,\ ]$ = the ith row;

$A[\ ,\mathscr{J}]$ = the submatrix of columns in $\mathscr{J}$; $A[\ ,j]$ = the jth column;

$A[\mathscr{I},\mathscr{J}]$ = the intersection submatrix; $A[i,j]$ = the (i,j)th element.

Similarly, for $\mathbf{x} \in R^n$:

$\mathbf{x}[\mathscr{J}]$ = the subvector indexed by $\mathscr{J}$; $\mathbf{x}[i]$ = the ith component.

For any subset $\mathscr{J}$ of $\overline{1,n}$:

$\#\mathscr{J}$ = number of elements of $\mathscr{J}$,
$\mathscr{J}^c$ = the complement of $\mathscr{J}$ (in $\overline{1,n}$).

For any two sets $\mathscr{J}$, $\mathscr{K}$:

$$\mathscr{J} \setminus \mathscr{K} := \mathscr{J} \cap \mathscr{K}^c = \text{the set-theoretic difference.}$$

Particular vectors and matrices include:

$\mathbf{0}$ = the zero vector of appropriate dimension,
O = the zero matrix of appropriate dimensions,
I_k = the $k \times k$ identity matrix.

Finally, a finite set $\mathscr{Z}$ of vectors is said to be *represented* by a matrix Z if

$$\{\text{columns } Z\} = \mathscr{Z}.$$

1. INTRODUCTION

Given $A \in R_m^{m\times n}$, $\mathbf{b} \in R^m$, $\mathbf{c} \in R^n$, the *simplex algorithm* [12, 13] for solving the (*primal*) *linear program*

$$\max\{\mathbf{c}^T\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0}\} \tag{P}$$

proceeds, while improvement is possible, along edges (connecting adjacent vertices) of the polyhedron

$$P := \{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0}\}. \tag{1}$$

Edges do not figure prominently in algebraic descriptions of the algorithm, given usually in terms of vertices (represented by basic feasible solutions). Alternatively, the simplex algorithm can be described explicitly in terms of the edges of P. At each iteration, a suitable basis of $N(A)$,

$$\mathscr{Z} = \{\mathbf{z}^1, \mathbf{z}^2, \dots, \mathbf{z}^{n-m}\}, \tag{2}$$

can be read from the simplex table,[1] giving the edges of P at the given basic feasible solution (b.f.s.) $\mathbf{x}$. The basis $\mathscr{Z}$ is called a *canonical basis* of $N(A)$.[2] The implementation of the simplex algorithm in terms of the canonical bases $\mathscr{Z}$ is called here the *$\mathscr{Z}$-simplex algorithm*. Its general iteration begins with a b.f.s. $\mathbf{x}$ and a canonical basis $\mathscr{Z}$ and proceeds as follows:

> **select** $\mathbf{z}^p \in \mathscr{Z}$ with $\mathbf{c}^T\mathbf{z}^p > 0$;
>   **if** none exists **then** $\mathbf{x}$ is optimal; **stop.**
> **let** $\theta_p := \max\{\theta \geqslant 0 : \mathbf{x} + \theta\mathbf{z}^p \geqslant \mathbf{0}\}$;
>   **if** $\theta_p = \infty$ then (P) is unbounded; **stop.**
> **update** $\mathbf{x} := \mathbf{x} + \theta_p\mathbf{z}^p$;
> **update** $\mathscr{Z}$.

Canonical bases were introduced by Orden [20] (in a different context) and Rockafellar ([22], where the vectors in $\mathscr{Z}$ are called *elementary vectors*).

Suitable bases $\mathscr{Z}$ of $N(A)$ were used in algorithms for linearly constrained optimization by Murtagh and Saunders [18], Gill et al. [14], and others. See also [6] and [11] for the computation of sparse bases.

The explicit use of edges in simplex algorithms allows an easy implementation of various criteria for edge selection [15], as well as new algorithms using nonnegative combinations of edges [16] and recursive programming [3].

Let us turn now to the *dual problem* of (P),

$$\min\{\mathbf{b}^T\mathbf{u} : A^T\mathbf{u} \geqslant \mathbf{c}\} \tag{D}$$

---

[1] In the sense that the information needed to construct $\mathscr{Z}$ is available in the simplex table.

[2] In the linear programming literature, especially in discussions of the simplex algorithm, "basis" usually means a basis of $R(A)$. The bases in this paper are of $N(A)$ and $R(A^T)$. The adjective "canonical" should prevent confusing them with the "simplex bases".

which, by changing variables,

$$\mathbf{y} = A^T\mathbf{u}, \tag{3}$$

can be rewritten as

$$\min\{\tilde{\mathbf{x}}^T\mathbf{y} : \mathbf{y} \in R(A^T), \mathbf{y} \geqslant \mathbf{c}\}, \tag{$\tilde{\text{D}}$}$$

where $\tilde{\mathbf{x}}$ is *any* solution of $A\mathbf{x} = \mathbf{b}$. The feasible set for the problem ($\tilde{\text{D}}$) is the polyhedron

$$\tilde{D} := \{\mathbf{y} : \mathbf{y} \in R(A^T), \mathbf{y} \geqslant \mathbf{c}\}, \tag{4}$$

and again, the dual method [17] for solving (D) can be described explicitly in terms of the edges of the polyhedron $\tilde{D}$. At each iteration, a canonical basis of $R(A^T)$,

$$\mathscr{Y} = \{\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^m\}, \tag{5}$$

can be read from the simplex table, giving the edges of $\tilde{D}$ at the given b.f.s. $\mathbf{y}$. The implementation of the dual method in terms of the canonical bases $\mathscr{Y}$ is called here the *$\mathscr{Y}$-dual method*. Its general iteration is as follows:

select $\mathbf{y}^p \in \mathscr{Y}$ with $\tilde{\mathbf{x}}^T\mathbf{y}^p < 0$;
  if none exists **then** $\mathbf{y}$ is optimal; **stop**.
let $\theta_p := \max\{\theta \geqslant 0 : \mathbf{y} + \theta\mathbf{y}^p \geqslant \mathbf{c}\}$;
  if $\theta_p = \infty$ then ($\tilde{\text{D}}$) is unbounded; **stop**.
update $\mathbf{y} := \mathbf{y} + \theta_p\mathbf{y}^p$;
update $\mathscr{Y}$.

Comparing this with the simplex iteration shows that both algorithms are essentially the same,[3] allowing a simple and unified development of LP, theory and algorithms. Besides its tutorial value, our approach has several practical advantages.

One advantage of canonical bases of $N(A)$ and $R(A^T)$ is that they undergo very simple changes when a new row is adjoined to, or when a row is deleted from, the matrix A. This observation is used in [3] to solve linear programs recursively, starting with a single equation of $A\mathbf{x} = \mathbf{b}$ and adjoining

[3] Both are feasible-direction methods [25], using extreme feasible directions which are the edges of the polyhedron in question at the given b.f.s.

equations as needed. At each iteration the optimal solution, and the canonical basis, are updated to accommodate the adjoined equation. The recursive (primal and dual) algorithms of [3] proceed along points which are not b.f.s.'s of (P) or ($\tilde{\mathrm{D}}$). The "priming" of the recursive algorithms of [3] is especially easy, since for a single equation, the optimal solution of (P) is immediate, saving the effort required by phase I of simplex algorithms.

Another advantage is that, by Corollary 1 below, one can switch easily (without computations) between the $\mathscr{X}$-simplex algorithm and the $\mathscr{Y}$-dual method. This flexibility is useful in cutting-plane algorithms for integer programming, and in postoptimal analysis. It may also save some work in phase I of simplex algorithms: In [5] canonical bases were used in the computation of a b.f.s. of (P) or of ($\tilde{\mathrm{D}}$), whichever comes first. If a b.f.s. of (P) is found, one continues with the $\mathscr{X}$-simplex algorithm. If a b.f.s. of ($\tilde{\mathrm{D}}$) turns out first, continuation is by the $\mathscr{Y}$-dual method.

The purpose of this paper is to give a unified development of simplex algorithms (primal and dual) with calculations in terms of canonical bases.

Relevant results about canonical bases are collected in Section 2. Theorem 1 gives matrix representations for the canonical bases $\mathscr{X}$ and $\mathscr{Y}$. Corollary 1 shows that either one of these bases is obtainable from the other without computations. Theorem 2 establishes a 1-1 correspondence between canonical bases and basic sets $\mathscr{J}$.

In Section 3 we present the computation of canonical bases using the Gauss-Jordan elimination, and the updating of canonical bases with changes in the underlying basic set $\mathscr{J}$. If two basic sets $\mathscr{J}$ and $\bar{\mathscr{J}}$ differ in exactly one vector, they are called *adjacent*, and the corresponding updating procedures,

ADJACENT $N(A)$ BASIS

and

ADJACENT $R(A^T)$ BASIS,

are the updates used in compact simplex versions. Theorem 3 gives the formula for general (nonadjacent) updates.

The geometric interpretation of canonical bases is given in Section 4. Theorem 4 shows that the vectors of $\mathscr{X}$ define edges, zero edges, or extreme rays of the polyhedron P. Theorem 5 does the same for the basis $\mathscr{Y}$ and the polyhedron $\tilde{D}$.

In Section 5 we present the $\mathscr{X}$-simplex algorithm and the $\mathscr{Y}$-dual method in their simplest form, not necessarily the form in which they should be implemented.

APL programs of these algorithms, and those of [3] and [5], are available from the author.

2. CANONICAL BASES OF $N(A)$ AND $R(A^T)$

In linear algebra, bases of subspaces are unordered sets. For ease of exposition and notational simplicity, especially in updating procedures, we define canonical bases as ordered sets.

Let $A \in R_m^{m \times n}$ and $\mathscr{J} \subset \overline{1,n}$ with $\#\mathscr{J} = m$ be fixed. The bases in the title are now defined.

DEFINITION 1.

(a) A basis $\mathscr{Z} = \{\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^{n-m}\}$ of $N(A)$ is *$\mathscr{J}$-canonical* if for each $j \in \overline{1, n-m}$ there is a $\zeta_j \in \mathscr{J}^c$ such that

$$\mathbf{z}^j[\zeta_j] = 1, \quad \mathbf{z}^k[\zeta_j] = 0 \qquad \text{for all} \quad j \neq k \in \overline{1, n-m}. \tag{6}$$

The ordered set $\zeta = \{\zeta_1, \zeta_2, \ldots, \zeta_{n-m}\}$, called the *list* of $\mathscr{Z}$, is part of the definition and is assumed given. Accordingly we denote a $\mathscr{J}$-canonical basis of $N(A)$ by $\{\mathscr{Z}, \zeta\}$, or just by $\mathscr{Z}$ if ζ is understood or is not needed. For a given $\mathscr{J}$ there are $(n-m)!$ different lists ζ, corresponding to the permutations of $\mathscr{J}^c$.

(b) A basis $\mathscr{Y} = \{\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^m\}$ of $R(A^T)$ is *$\mathscr{J}$-canonical* if for each $j \in \overline{1, m}$ there is a $v_j \in \mathscr{J}$ such that

$$\mathbf{y}^j[v_j] = 1, \quad \mathbf{y}^k[v_j] = 0 \qquad \text{for all} \quad j \neq k \in \overline{1, m}. \tag{7}$$

The ordered set $v = \{v_1, v_2, \ldots, v_m\}$ is called the *list* of $\mathscr{Y}$, and is part of the definition. There are $m!$ different lists v of $\mathscr{J}$, and a $\mathscr{J}$-canonical basis of $R(A^T)$ is denoted by $\{\mathscr{Y}, v\}$ or just by $\mathscr{Y}$.

An immediate consequence is:

LEMMA 1. *Given a set $\mathscr{J}$ and a list, if a $\mathscr{J}$-canonical basis exists, it is unique.*

Proof. By definition, a canonical basis is an ordered set in which every vector has component 1 where all other vectors have zeros. Such a basis is unique. ∎

From a $\mathscr{J}$-canonical basis of $N(A)$, one can get a $\mathscr{J}$-canonical basis of $R(A^T)$, and conversely.

THEOREM 1. *Let $\mathscr{J} \subset \overline{1,n}$, $\#\mathscr{J} = m$, let ζ and v be lists of $\mathscr{J}^c$ and $\mathscr{J}$ respectively, and let M be the $n \times n$ permutation matrix mapping*

$$\overline{1,m} \quad \text{to} \quad \{v_1, v_2, \dots, v_m\},$$

and

$$\overline{m+1,n} \quad \text{to} \quad \{\zeta_1, \zeta_2, \dots, \zeta_{n-m}\},$$

i.e.

$$M[v_i, i] = 1 \qquad \text{for} \quad i \in \overline{1,m},$$

$$M[\zeta_i, m+i] = 1 \qquad \text{for} \quad i \in \overline{1, n-m},$$

and all other $M[i,j] = 0$.

Let $\{\mathscr{X}, \zeta\}$ be a $\mathscr{J}$-canonical basis of $N(A)$. Then $\mathscr{X}$ is represented by the matrix

$$Z = M\begin{pmatrix} -X \\ I_{n-m} \end{pmatrix} \tag{8}$$

for some $X \in R^{m \times (n-m)}$, in which case $\{\mathscr{Y}, v\}$ is a $\mathscr{J}$-canonical basis of $R(A^T)$, with $\mathscr{Y}$ represented by

$$Y = M\begin{pmatrix} I_m \\ X^T \end{pmatrix} \tag{9}$$

Conversely, if the $\mathscr{J}$-canonical basis of $R(A^T)$ is represented by (9), *then the $\mathscr{J}$-canonical basis of $N(A)$ is represented by* (8).

Proof. If $\{\mathscr{X}, \zeta\}$ is a $\mathscr{J}$-canonical basis of $N(A)$ then, by Definition 1(a), $\mathscr{X}$ is represented by the matrix (8) whose uniqueness follows from Lemma 1. Since the columns of the matrix (9) are orthogonal to those of (8), it follows from Definition 1(b) [using the fact that $\{N(A), R(A^T)\}$ are complementary orthogonal subspaces] that the matrix (9) represents a $\mathscr{J}$-canonical basis of $R(A^T)$, with uniqueness guaranteed by Lemma 1. The converse statement is proved analogously. ∎

Since (8) and (9) use the same matrix X, one canonical basis is obtainable from the other by a *transposition* and *change of sign*. This situation is summarized in the following corollary, a special case of Rockafellar's general basis theorem, [23, p. 457].

COROLLARY 1. *A $\mathscr{J}$-canonical basis of $N(A)$ exists if and only if a $\mathscr{J}$-canonical basis of $R(A^T)$ exists, in which case one is obtainable from the other without computation. Specifically:*

(a) *If $\{\mathscr{Z}, \zeta\}$ is a $\mathscr{J}$-canonical basis of $N(A)$, and if v is any listing of the elements of $\mathscr{J}$, then $\{\mathscr{Y}, v\}$ is a $\mathscr{J}$-canonical basis of $R(A^T)$, where*

$$\mathscr{Y} = \{\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^m\}$$

is given by

$$\mathbf{y}^k[\zeta_j] = -\mathbf{z}^j[v_k] \qquad \text{for} \quad j \in \overline{1, n-m}, \quad k \in \overline{1, m}, \tag{10}$$

and the remaining elements of $\mathbf{y}^k$ (zeros and ones) specified by (7).

(b) *If $\{\mathscr{Y}, v\}$ is a $\mathscr{J}$-canonical basis of $R(A^T)$, and if ζ is any listing of the elements of $\mathscr{J}^c$, then $\{\mathscr{Z}, \zeta\}$ is a $\mathscr{J}$-canonical basis of $N(A)$, where*

$$\mathscr{Z} = \{\mathbf{z}^1, \mathbf{z}^2, \dots, \mathbf{z}^{n-m}\}$$

is given by (10) *and* (6).

The question "For which sets $\mathscr{J}$ do canonical bases exist?" is answered in Theorem 2 below. First we need:

DEFINITION 2. Given $A \in R_m^{m \times n}$, a subset $\mathscr{J} \subset \overline{1, n}$ with $\#\mathscr{J} = m$ is called *basic* if the matrix $A[\ ,\mathscr{J}]$ is nonsingular.

THEOREM 2. *$\mathscr{J}$-canonical bases $\{\mathscr{Z}, \mathscr{Y}\}$ of $\{N(A), R(A^T)\}$ exist if and only if $\mathscr{J}$ is basic, in which case $\mathscr{Z}$, $\mathscr{Y}$ are represented by* (8), (9) *with*

$$X = A[\ ,\mathscr{J}]^{-1} A[\ ,\mathscr{J}^c]. \tag{11}$$

Proof. If $A[\ ,\mathscr{J}]$ is nonsingular, then substituting (11) in (8), (9) gives $\mathscr{J}$-canonical bases of $N(A)$, $R(A^T)$.

Conversely, let $A[\ ,\mathscr{J}]$ be singular. Then any nonzero vector in $N(A[\ ,\mathscr{J}])$ can be padded with $n-m$ zeros to give a vector $\mathbf{0} \neq \mathbf{x} \in N(A)$

with $\mathbf{x}[\mathscr{J}^c] = \mathbf{0}$. By Definition 1(a), such a vector is not a linear combination of a $\mathscr{J}$-canonical basis of $N(A)$, contradicting the existence of $\mathscr{J}$-canonical bases. ∎

3. COMPUTATION AND UPDATING OF CANONICAL BASES

The computation of canonical bases using Gauss-Jordan elimination is straightforward; see e.g. [19, Theorems 4.8 and 5.4].

Given $A \in R_m^{m \times n}$ and a basic $\mathscr{J} \subset \overline{1,n}$, $\mathscr{J}$-canonical bases $\{\mathscr{X}, \mathscr{Y}\}$ of $\{N(A), R(A^T)\}$ respectively are computed as follows:

compute H the Hermite (row echelon) form of A
  using pivots only in $A[\ ,\mathscr{J}]$;
let $\mathscr{Y} = \{\text{rows}\, H\}$;
read $\mathscr{X}$ from $\mathscr{Y}$ by Corollary 1(b).

EXAMPLE 1.

$$A = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Here $\mathscr{J} = \{4,5,6\}$ is basic, and the Hermite form already given. Writing vectors as row vectors, it follows that

$$\mathbf{y}^1 = (2 \quad 1 \quad 1 \quad 1 \quad 0 \quad 0),$$

$$\mathbf{y}^2 = (1 \quad 2 \quad 3 \quad 0 \quad 1 \quad 0),$$

$$\mathbf{y}^3 = (2 \quad 2 \quad 1 \quad 0 \quad 0 \quad 1)$$

is the $\mathscr{J}$-canonical basis of $R(A^T)$ with $\{v_1, v_2, v_3\} = \{4,5,6\}$.

For the list $\{\zeta_1, \zeta_2, \zeta_3\} = \{1,2,3\}$ the $\mathscr{J}$-canonical basis of $N(A)$ is, by (10) and (6),

$$\mathbf{z}^1 = (1 \quad 0 \quad 0 \quad -2 \quad -1 \quad -2),$$

$$\mathbf{z}^2 = (0 \quad 1 \quad 0 \quad -1 \quad -2 \quad -2),$$

$$\mathbf{z}^3 = (0 \quad 0 \quad 1 \quad -1 \quad -3 \quad -1).$$

Let $\mathcal{J} \subset \overline{1,n}$ be basic, and let $\mathcal{Z}$ be the $\mathcal{J}$-canonical basis of $N(A)$. Let $\bar{\mathcal{J}}$ be any subset of $\overline{1,n}$ with

$$\#\bar{\mathcal{J}} = m, \qquad \bar{\mathcal{J}} \neq \mathcal{J}$$

Two natural questions are: What conditions in terms of $\{\mathcal{J}, \mathcal{Z}\}$ guarantee that $\bar{\mathcal{J}}$ is basic? And is so, can the $\bar{\mathcal{J}}$-canonical basis of $N(A)$ be given in terms of the known basis $\mathcal{Z}$? These questions are answered in Theorem 3 below, using the following notation, based on Definition 1(a): For any subset $\mathcal{K} \subset \mathcal{J}^c$

$$\zeta^{-1}(\mathcal{K}) := \{ j : \zeta_j \in \mathcal{K} \}. \tag{12}$$

THEOREM 3. *Let $\mathcal{J}$, $\mathcal{Z}$, $\bar{\mathcal{J}}$ be as above, and let the matrix Z represent $\mathcal{Z}$. Then $\bar{\mathcal{J}}$ is basic if and only if the submatrix*

$$T = \mathrm{Z}[\mathcal{J} \setminus \bar{\mathcal{J}}, \zeta^{-1}(\bar{\mathcal{J}} \setminus \mathcal{J})] \tag{13}$$

is nonsingular, in which case the $\bar{\mathcal{J}}$-canonical basis of $N(A)$, $\bar{\mathcal{Z}}$, is represented by the matrix $\bar{\mathrm{Z}}$ given as follows:

$$\bar{\mathrm{Z}}[\ , \zeta^{-1}(\bar{\mathcal{J}} \setminus \mathcal{J})] = \mathrm{Z}[\ , \zeta^{-1}(\bar{\mathcal{J}} \setminus \mathcal{J})] T^{-1}, \tag{14}$$

$$\begin{aligned} \bar{\mathrm{Z}}[\ , \zeta^{-1}(\mathcal{J}^c \setminus \bar{\mathcal{J}})] &= \mathrm{Z}[\ , \zeta^{-1}(\mathcal{J}^c \setminus \bar{\mathcal{J}})] \\ &\quad - \bar{\mathrm{Z}}[\ , \zeta^{-1}(\bar{\mathcal{J}} \setminus \mathcal{J})] \mathrm{Z}[\mathcal{J} \setminus \bar{\mathcal{J}}, \zeta^{-1}(\mathcal{J}^c \setminus \bar{\mathcal{J}})]. \end{aligned} \tag{15}$$

Proof. "If": Let T be nonsingular. Then it can be verified using Definition 1(a) that $\bar{\mathcal{Z}}$, given by (14) and (15), is a $\bar{\mathcal{J}}$-canonical basis of $N(A)$.

"Only if": Let T be singular. From Definition 1(a) we have, for any $\bar{\mathcal{J}}$,

$$\mathrm{Z}[\mathcal{J}^c \setminus \bar{\mathcal{J}}, \zeta^{-1}(\bar{\mathcal{J}} \setminus \mathcal{J})] = O.$$

Thus the singularity of T and the identity

$$\bar{\mathcal{J}}^c = (\mathcal{J} \setminus \bar{\mathcal{J}}) \cup (\mathcal{J}^c \setminus \bar{\mathcal{J}})$$

together imply that the columns of

$$Z[\bar{\mathscr{J}}^c, \zeta^{-1}(\bar{\mathscr{J}} \setminus \mathscr{J})]$$

are linearly dependent. Therefore, the matrix $Z[\bar{\mathscr{J}}^c, \ ]$ is singular, and so is its transpose $Z^T[\ , \bar{\mathscr{J}}^c]$. Any nonzero vector in $N(Z^T[\ , \bar{\mathscr{J}}^c])$ can be padded with zeros in $\bar{\mathscr{J}}$ to give a nonzero vector $\mathbf{x}$ such that

$$\mathbf{x} \in N(Z^T), \tag{16}$$

$$\mathbf{x}[\bar{\mathscr{J}}] = \mathbf{0} \tag{17}$$

Now (16) means that $\mathbf{x}$ is perpendicular to $N(A)$; hence $\mathbf{x} \in R(A^T)$. But by (17) and Definition 1(b), $\mathbf{x}$ is not a linear combination of a $\bar{\mathscr{J}}$-canonical basis of $R(A^T)$, contradicting the existence of such a basis. Therefore $\bar{\mathscr{J}}$ is not basic. ∎

If $\mathscr{J}$ and $\bar{\mathscr{J}}$ differ by two or more vectors, Theorem 3 describes a pivoting operation with a matrix pivot. The same nonsingularity condition appears in studies of simplex updates by Bisschop and Meeraus [7] and Gill et al. [14].

Theorem 3 can alternatively be stated in terms of the canonical bases of $R(A^T)$, using Corollary 1. Thus, $\bar{\mathscr{J}}$ is basic if and only if the submatrix

$$S = Y[\bar{\mathscr{J}} \setminus \mathscr{J}, v^{-1}(\mathscr{J} \setminus \bar{\mathscr{J}})] \tag{18}$$

is nonsingular, where the matrix Y represents the $\mathscr{J}$-canonical basis $\mathscr{Y}$ of $R(A^T)$, and where, for any $\mathscr{K} \subset \mathscr{J}$,

$$v^{-1}(\mathscr{K}) := \{ j : v_j \in \mathscr{K} \}; \tag{19}$$

see Definition 1(b). The $\bar{\mathscr{J}}$-canonical basis of $R(A^T)$ is then computed analogously to (14), (15).

Example 2. Let A, $\mathscr{J}$ be as in Example 1, and let $\bar{\mathscr{J}} = \{1,3,6\}$. Here

$$Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -1 & -1 \\ -1 & -2 & -3 \\ -2 & -2 & -1 \end{pmatrix},$$

and from (13),

$$T = Z[\mathscr{I} \setminus \bar{\mathscr{I}}, \zeta^{-1}(\bar{\mathscr{I}} \setminus \mathscr{I})] = Z[\{4,5\},\{1,3\}] = \begin{pmatrix} -2 & -1 \\ -1 & -3 \end{pmatrix}$$

is nonsingular. Therefore, $\bar{\mathscr{I}}$ is basic, and by (14), (15),

$$\bar{Z}[\ ,\zeta^{-1}(\bar{\mathscr{I}} \setminus \mathscr{I})] = \bar{Z}[\ ,\{1,3\}] = Z[\ ,\{1,3\}]T^{-1}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ -2 & -1 \\ -1 & -3 \\ -2 & -1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} -3 & 1 \\ 1 & -2 \end{pmatrix}$$

$$= \frac{1}{5} \begin{pmatrix} -3 & 1 \\ 0 & 0 \\ 1 & -2 \\ 5 & 0 \\ 0 & 5 \\ 5 & 0 \end{pmatrix},$$

$$\bar{Z}[\ ,\zeta^{-1}(\mathscr{I}^c \setminus \bar{\mathscr{I}})] = \bar{Z}[\ ,2] = Z[\ ,2] - \bar{Z}[\ ,\{1,3\}]Z[\{4,5\},2]$$

$$= \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ -2 \\ -2 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} -3 & 1 \\ 0 & 0 \\ 1 & -2 \\ 5 & 0 \\ 0 & 5 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

$$= \frac{1}{5} \begin{pmatrix} -1 \\ 5 \\ -3 \\ 0 \\ 0 \\ -5 \end{pmatrix}.$$

The $\bar{\mathscr{I}}$-canonical basis of $N(A)$ is therefore represented by

$$\bar{Z} = \begin{pmatrix} -\frac{3}{5} & -\frac{1}{5} & \frac{1}{5} \\ 0 & 1 & 0 \\ \frac{1}{5} & -\frac{3}{5} & -\frac{2}{5} \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}, \qquad \text{with} \quad \{\bar{\zeta}_1, \bar{\zeta}_2, \bar{\zeta}_3\} = \{4,2,5\}$$

By Corollary 1, the $\bar{\mathscr{I}}$-canonical basis of $R(A^T)$ is represented by

$$\bar{Y} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{5} & \frac{3}{5} & 1 \\ 0 & 1 & 0 \\ \frac{3}{5} & -\frac{1}{5} & -1 \\ -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \text{with} \quad \{\bar{v}_1, \bar{v}_2, \bar{v}_3\} = \{1,3,6\}.$$

Theorem 3 is useful if the "pivot" matrix T in (13) [or S in (18)] is small-dimensional, i.e. if there is a considerable overlap between $\mathscr{I}, \bar{\mathscr{I}}$.

The least favorable case is when $\mathscr{I}, \bar{\mathscr{I}}$ are disjoint (possible only if $m \leqslant n/2$). Here Theorem 3 offers no advantage over computation from scratch.

The most favorable case for applying Theorem 3 is when

$$\#(\mathscr{I} \cap \bar{\mathscr{I}}) = m - 1, \tag{20}$$

in which case we call the basic sets $\{\mathscr{I}, \bar{\mathscr{I}}\}$, and the corresponding canonical bases, *adjacent*. Here the pivot matrices (13) and (18) are 1×1, and (14), (15) simplify to the "pivoting" used in simplex-type algorithms.

For later use we give now the updating procedures for *adjacent canonical bases*, using Pidgin ALGOL as in [21]. First, the procedure for updating the canonical basis of $N(A)$.

Procedure ADJACENT $N(A)$ BASIS:

input: A basic set $\mathscr{I}$;

$\mathscr{Z} = \{\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^{n-m}\}$ the $\mathscr{I}$-canonical basis of $N(A)$, and its list $\zeta = \{\zeta_1, \zeta_2, \ldots, \zeta_{n-m}\}$;

A set $\bar{\mathscr{I}} \subset \overline{1,n}$ with

(i) $\#(\mathscr{I} \cap \bar{\mathscr{I}}) = m - 1$,

(ii) $\mathbf{z}^p[q] \neq 0$, where $\{q\} = \mathscr{I} \setminus \bar{\mathscr{I}}$, $\{p\} = \zeta^{-1}(\bar{\mathscr{I}} \setminus \mathscr{I})$.

output: $\bar{\mathscr{Z}} = \{\bar{\mathbf{z}}^1, \bar{\mathbf{z}}^2, \ldots, \bar{\mathbf{z}}^{n-m}\}$ the $\bar{\mathscr{J}}$-canonical basis of $N(A)$, and its list $\bar{\zeta} = \{\bar{\zeta}_1, \bar{\zeta}_2, \ldots, \bar{\zeta}_{n-m}\}$.

begin

$$\bar{\mathbf{z}}^p := \frac{\mathbf{z}^p}{\mathbf{z}^p[q]};$$

$$\bar{\zeta}_p := q;$$

while $j = 1, \ldots, n-m$ **and** $j \neq p$ **do**

$$\bar{\mathbf{z}}^j := \mathbf{z}^j - \mathbf{z}^j[q]\bar{\mathbf{z}}^p;$$

$$\bar{\zeta}_j := \zeta_j;$$

end.

The procedure for updating the canonical basis of $R(A^T)$ is:

Procedure ADJACENT $R(A^T)$ BASIS:

input: A basic set $\mathscr{J}$;
$\mathscr{Y} = \{\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^m\}$ the $\mathscr{J}$-canonical basis of $R(A^T)$, and its list $\upsilon = \{\upsilon_1, \upsilon_2, \ldots, \upsilon_m\}$;
A set $\bar{\mathscr{J}} \subset \overline{1,n}$ with
(i) $\#(\mathscr{J} \cap \bar{\mathscr{J}}) = m - 1$
(ii) $\mathbf{y}^p[q] \neq 0$ where $\{q\} = \bar{\mathscr{J}} \setminus \mathscr{J}$, $\{p\} = \upsilon^{-1}(\mathscr{J} \setminus \bar{\mathscr{J}})$.

output: $\bar{\mathscr{Y}} = \{\bar{\mathbf{y}}^1, \bar{\mathbf{y}}^2, \ldots, \bar{\mathbf{y}}^m\}$ the $\bar{\mathscr{J}}$-canonical basis of $R(A^T)$, and its list $\bar{\upsilon} = \{\bar{\upsilon}_1, \bar{\upsilon}_2, \ldots, \bar{\upsilon}_m\}$.

begin

$$\bar{\mathbf{y}}^p := \frac{\mathbf{y}^p}{\mathbf{y}^p[q]};$$

$$\bar{\upsilon}^p := q;$$

while $j = 1, \ldots, m$ **and** $j \neq p$ **do**

$$\bar{\mathbf{y}}^j := \mathbf{y}^j - \mathbf{y}^j[q]\bar{\mathbf{y}}^p;$$

$$\bar{\upsilon}_j := \upsilon_j;$$

end.

By Corollary 1 only one of the bases $\{\mathscr{X}, \mathscr{Y}\}$ needs updating, so only one of the above procedures is needed when passing from $\mathscr{J}$ to an adjacent $\bar{\mathscr{J}}$.

4. CANONICAL BASES AND CONVEX POLYHEDRA

For given $A \in R_m^{m \times n}$, $\mathbf{b} \in R^m$, $\mathbf{c} \in R^n$ consider the dual pair of linear programs:

$$\max\{\mathbf{c}^T\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0}\} \qquad (\text{the } \textit{primal problem}), \tag{P}$$

$$\min\{\mathbf{b}^T\mathbf{u} : A^T\mathbf{u} \geqslant \mathbf{c}\} \qquad (\text{the } \textit{dual}). \tag{D}$$

For our purposes it is convenient to rewrite (D) as a problem in the same space with (P). Changing the variables of (D) from $\mathbf{u}$ to $\mathbf{y}$ according to (3), it follows (since A is of full row rank) that the correspondence

$$\{\mathbf{u} \in R^m\} \mapsto \{\mathbf{y} \in R(A^T)\}$$

is 1-1 and onto. For *any* solution $\tilde{\mathbf{x}}$ of the linear equation

$$A\mathbf{x} = \mathbf{b} \tag{21}$$

the dual problem (D) is equivalent to the problem

$$\min\{\tilde{\mathbf{x}}^T\mathbf{y} : \mathbf{y} \in R(A^T), \mathbf{y} \geqslant \mathbf{c}\}. \tag{$\tilde{\text{D}}$}$$

Indeed, by (3) and (21),

$$\tilde{\mathbf{x}}^T\mathbf{y} = \tilde{\mathbf{x}}^T A^T\mathbf{u} = \mathbf{b}^T\mathbf{u}$$

independently of the $\tilde{\mathbf{x}}$ used in ($\tilde{\text{D}}$). The economic interpretation of the new variables $\mathbf{y}$ (at optimum) is that of *activation prices*, i.e. minimal prices at which unused activities become competitive (with all other prices unchanged); see [4].

Given a basic index set $\mathscr{J} \subset \overline{1, n}$, a point $\mathbf{x}$ is the *$\mathscr{J}$-basic feasible solution* (abbreviated $\mathscr{J}$-b.f.s. or just b.f.s.) of (P) if it satisfies (21),

$$\mathbf{x} \geqslant \mathbf{0}, \tag{22}$$

and

$$\mathbf{x}[\mathscr{J}^c] = \mathbf{0}. \tag{23}$$

$\mathbf{x}$ is a *feasible solution* of (P) if (21) and (22) hold, and the *$\mathscr{J}$-basic solution* of (P) if (21) and (23) hold. The $\mathscr{J}$-basic solution $\mathbf{x}$ is given by

$$\mathbf{x}[\mathscr{J}^c] = \mathbf{0}, \qquad \mathbf{x}[\mathscr{J}] = A[\ ,\mathscr{J}]^{-1}\mathbf{b}. \tag{24}$$

A $\mathscr{J}$-b.f.s. $\mathbf{x}$ is *nondegenerate* if

$$\mathbf{x}[j] > 0 \qquad \text{for all} \quad j \in \mathscr{J} \tag{25}$$

and is *degenerate* otherwise.

The *set of feasible solutions* of (P) is the polyhedron P of (1). Recall that a b.f.s. of (P) is a *vertex* of P.

The following theorem shows that at any basic solution, the feasible polyhedron P is contained in the cone spanned by the corresponding canonical basis $\mathscr{Z}$ of $N(A)$. If the given basic solution $\mathbf{x}^0$ is feasible, then the vectors of $\mathscr{Z}$ are (directions of) edges, zero edges, or extreme rays of P at $\mathbf{x}^0$.

THEOREM 4. *Let $\mathscr{J}$ be a basic index set, and let $\mathscr{Z} = \{\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^{n-m}\}$ be the $\mathscr{J}$-canonical basis of $N(A)$.*

(a) *If $\mathbf{x}^0$ is the $\mathscr{J}$-basic solution of (P) (not necessarily feasible), then*

$$P \subset \mathbf{x}^0 + \left\{ \sum_{j=1}^{n-m} \lambda_j \mathbf{z}^j : \lambda_j \geqslant 0,\, j \in \overline{1, n-m} \right\}. \tag{26}$$

(b) *If $\mathbf{x}^0$ is also feasible (i.e. $\mathbf{x}^0$ is the $\mathscr{J}$-b.f.s. of (P)), let for $j \in \overline{1, n-m}$*

$$\theta_j := \begin{cases} \min\left\{ -\dfrac{\mathbf{x}^0[i]}{\mathbf{z}^j[i]} : i \in \mathscr{J}, \mathbf{z}^j[i] < 0 \right\}, \\ \infty \text{ if } \mathbf{z}^j[i] \geqslant 0 \text{ for all } i \in \mathscr{J}. \end{cases} \tag{27}$$

Then, for $j \in \overline{1, n-m}$,

(b1) *$\mathbf{x}^0 + \lambda \mathbf{z}^j \in P$ for all $0 \leqslant \lambda < \theta_j$.*

(b2) *If $\theta_j < \infty$ then*

$$\bar{\mathbf{x}} := \mathbf{x}^0 + \theta_j \mathbf{z}^j \tag{28}$$

is the $\bar{\mathscr{J}}$*-b. f. s. of* (P), *where*

$$\bar{\mathscr{J}} = (\mathscr{J} \setminus \{k\}) \cup \{\zeta_j\} \tag{29}$$

and k is (chosen from among)

$$k \in \arg\min\left\{ -\frac{\mathbf{x}^0[i]}{\mathbf{z}^j[i]} : i \in \mathscr{J}, \mathbf{z}^j[i] < 0 \right\}. \tag{30}$$

Proof. (a): Clearly

$$P \subset \{\mathbf{x}: A\mathbf{x} = \mathbf{b}\} = \mathbf{x}^0 + \left\{ \sum_{j=1}^{n-m} \lambda_j \mathbf{z}^j : \lambda_j \text{ real},\ j \in \overline{1, n-m} \right\},$$

but any $\lambda_j < 0$ will result in $\mathbf{x}[\zeta_j] < 0$ by Definition 1(a) and (23).

(b): Here we look for the maximal θ_j with $\mathbf{x}^0 + \theta_j \mathbf{z}^j \geqslant \mathbf{0}$. Since $\mathbf{x}^0$ is feasible, $\theta_j \geqslant 0$ and is given by (27). Conclusion (b1) is obvious.

If $\theta_j < \infty$, then by (27) and (28)

$$A\bar{\mathbf{x}} = \mathbf{b}, \qquad \bar{\mathbf{x}} \geqslant \mathbf{0}$$

Also, $\bar{\mathbf{x}}[k] = 0$ by (30), so from (29),

$$\bar{\mathbf{x}}[\bar{\mathscr{J}}^c] = \mathbf{0}.$$

Now the set $\bar{\mathscr{J}}$ is basic, by Theorem 3, since

$$\mathbf{z}^j[k] \neq 0, \qquad \text{where} \quad \{j\} = \zeta^{-1}(\bar{\mathscr{J}} \setminus \mathscr{J}), \quad \{k\} = \mathscr{J} \setminus \bar{\mathscr{J}},$$

and therefore $\bar{\mathbf{x}}$ is the $\bar{\mathscr{J}}$-b.f.s. of (P). ∎

If $\theta_j = \infty$ then, from (b1),

$$\mathbf{x}^0 + \lambda \mathbf{z}^j \in P \qquad \text{for all } 0 \leqslant \lambda$$

and $\mathbf{x}^0 + \{\lambda \mathbf{z}^j : \lambda \geqslant 0\}$ is an *extreme ray* of P, with $\mathbf{z}^j$ its *direction*; see e.g. Bazaraa and Jarvis [1, Appendix, Lemma 2].

If $0 < \theta_j < \infty$ then

$$\mathbf{x}^0 + \left\{ \lambda \mathbf{z}^j : 0 \leqslant \lambda \leqslant \theta_j \right\} \tag{31}$$

is the *edge* of P leading from the vertex $\mathbf{x}^0$ [the $\mathscr{J}$-b.f.s. of (P)] to the adjacent vertex $\bar{\mathbf{x}}$, the $\bar{\mathscr{J}}$-b.f.s.

If $\theta_j = 0$ we call (31) a *zero edge*.

Each b.f.s. of (P) thus corresponds to a canonical basis of $N(A)$ whose $n - m$ vectors define extreme rays, edges, or zero edges of P. If a b.f.s. is nondegenerate then it does not have zero edges. The converse is not true, as shown by the following example (see also the discussion in Chvátal [10, pp. 258–259]).

EXAMPLE 3. Let

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Then the basic sets $\mathscr{J}^1 = \{1,2,3\}$ and $\mathscr{J}^2 = \{1,2,4\}$ have the same (degenerate) b.f.s.

$$\mathbf{x} = (1 \quad 1 \quad 0 \quad 0 \quad 0)$$

and the same canonical basis of $N(A)$,

$$\mathbf{z}^1 = (-1 \quad -1 \quad 1 \quad 1 \quad 0) \qquad \text{with} \quad \theta_1 = 1,$$

$$\mathbf{z}^2 = (0 \quad -1 \quad 0 \quad 0 \quad 1) \qquad \text{with} \quad \theta_2 = 1.$$

At least $n - m$ edges or rays meet at each vertex of P. If the vertex is nondegenerate, then exactly $n - m$ edges or rays (but no zero edges) meet at it. Example 3 shows the converse to be false.

Consider now the dual problem ($\tilde{\mathrm{D}}$). Given a basic set $\mathscr{J} \subset \overline{1, n}$, a point $\mathbf{y}$ is the *$\mathscr{J}$-basic feasible solution* of ($\tilde{\mathrm{D}}$) if

$$\mathbf{y} \in R(A^T), \tag{32}$$

$$\mathbf{y} \geqslant \mathbf{c}, \tag{33}$$

$$\mathbf{y}[\mathscr{J}] = c[\mathscr{J}]. \tag{34}$$

$\mathbf{y}$ is a *feasible solution* of ($\tilde{\mathrm{D}}$) if (32) and (33) hold, and the *$\mathscr{J}$-basic solution* of ($\tilde{\mathrm{D}}$) if (32) and (34) hold. The $\mathscr{J}$-basic solution $\mathbf{y}$ is given by

$$\mathbf{y}^T = \mathbf{c}[\mathscr{J}]^T A[\ ,\mathscr{J}]^{-1} A. \tag{35}$$

The $\mathcal{J}$-b.f.s. y is *nondegenerate* if

$$\mathrm{y}[j] > \mathrm{c}[j] \qquad \text{for all} \quad j \in \mathcal{J}^c \tag{36}$$

and is *degenerate* otherwise.

The *set of feasible solutions* of ($\tilde{\mathrm{D}}$) is the polyhedron $\tilde{D}$ of (4).

The following theorem is an analogue of Theorem 4.

THEOREM 5. *Let $\mathcal{J}$ be a basic index set, and let $\mathcal{Y} = \{\mathrm{y}^1, \mathrm{y}^2, \dots, \mathrm{y}^m\}$ be the $\mathcal{J}$-canonical basis of $R(A^T)$.*

(a) *If y^0 is the $\mathcal{J}$-basic solution of* ($\tilde{\mathrm{D}}$), *then*

$$\tilde{D} \subset \mathrm{y}^0 + \left\{ \sum_{j=1}^{m} \lambda_j \mathrm{y}^j : \lambda_j \geqslant 0,\, j \in \overline{1,m} \right\}. \tag{37}$$

(b) *If y^0 is also feasible, let for $j \in \overline{1,m}$*

$$\theta_j := \begin{cases} \min\left\{ -\dfrac{\mathrm{y}^0[i] - \mathrm{c}[i]}{\mathrm{y}^j[i]} : i \in \mathcal{J}^c, \mathrm{y}^j[i] < 0 \right\}, \\ \infty \text{ if } \mathrm{y}^j[i] \geqslant 0 \text{ for all } i \in \mathcal{J}^c. \end{cases} \tag{38}$$

Then, for $j \in \overline{1,m}$,

(b1) $\mathrm{y}^0 + \lambda \mathrm{y}^j \in \tilde{D}$ *for all* $0 \leqslant \lambda < \theta_j$.

(b2) *If $\theta_j < \infty$, then*

$$\bar{\mathrm{y}} := \mathrm{y}^0 + \theta_j \mathrm{y}^j \tag{39}$$

is the $\bar{\mathcal{J}}$-b.f.s. of ($\tilde{\mathrm{D}}$), where

$$\bar{\mathcal{J}} = (\mathcal{J} \setminus \{v_j\}) \cup \{k\} \tag{40}$$

and

$$k \in \arg\min\left\{ -\frac{\mathrm{y}^0[i] - \mathrm{c}[i]}{\mathrm{y}^j[i]} : i \in \mathcal{J}^c, \mathrm{y}^j[i] < 0 \right\}. \tag{41}$$

Again, if the b.f.s. $\mathbf{y}^0$ is degenerate, it is possible for $\theta_j = 0$, resulting in $\mathbf{y}^0 = \bar{\mathbf{y}}$.

5. PRIMAL AND DUAL SIMPLEX ALGORITHMS USING CANONICAL BASES

The simplex algorithm [12] proceeds along edges of P, given by canonical bases $\mathscr{X}$ of $N(A)$. It follows that the simplex algorithm can be described (and implemented) in terms of these bases. Such an algorithm, presented below, is called a $\mathscr{X}$-simplex algorithm.

The dual method [17] can similarly be described in terms of the edges of $\tilde{D}$, that is canonical bases $\mathscr{Y}$ of $R(A^T)$. We call such an algorithm a *$\mathscr{Y}$-dual method*.

Both the $\mathscr{X}$-simplex and $\mathscr{Y}$-dual algorithms use inner products to compute the *reduced costs*:

LEMMA 2. *Let $\mathscr{J}$ be a basic set, let $\{\mathscr{X}, \zeta\}$ and $\{\mathscr{Y}, v\}$ be the $\mathscr{J}$-canonical bases of $\{N(A), R(A^T)\}$, let $\{\mathbf{x},\mathbf{y}\}$ be the $\mathscr{J}$-basic solutions (not necessarily feasible) of $\{(P),(\tilde{D})\}$, and let $\tilde{\mathbf{x}}$ be any solution of $A\mathbf{x} = \mathbf{b}$. Then*

$$-\mathbf{c}^T\mathbf{z}^j = \mathbf{y}[\zeta_j] - \mathbf{c}[\zeta_j] \qquad \text{for} \quad j \in \overline{1, n-m}, \tag{42}$$

$$\tilde{\mathbf{x}}^T\mathbf{y}^j = \mathbf{x}[v_j] \qquad \text{for} \quad j \in \overline{1, m}. \tag{43}$$

Proof. (42): From (34) and Definition 1(b), for any $j \in \overline{1, n-m}$,

$$\begin{aligned}(\mathbf{y} - \mathbf{c})^T\mathbf{z}^j &= (\mathbf{y}[\mathscr{J}^c] - \mathbf{c}[\mathscr{J}^c])^T\mathbf{z}^j[\mathscr{J}^c] \\ &= \mathbf{y}[\zeta_j] - \mathbf{c}[\zeta_j].\end{aligned} \tag{44}$$

Since $\mathbf{y} \in R(A^T)$, it follows that $\mathbf{y}^T\mathbf{z}^j = 0$, so that

$$(\mathbf{y} - \mathbf{c})^T\mathbf{z}^j = -\mathbf{c}^T\mathbf{z}^j,$$

which, together with (44), proves (42).

(43): Since $\tilde{\mathbf{x}} - \mathbf{x} \in N(A)$, it follows, for any $j \in \overline{1, m}$, that

$$\tilde{\mathbf{x}}^T \mathbf{y}^j = \mathbf{x}^T \mathbf{y}^j,$$

and (43) follows from (23) and Definition 1(b). ■

From (42) and (35) it follows that the inner products $\mathbf{c}^T \mathbf{z}^j$ equal the simplex reduced costs

$$\mathbf{c}^T \mathbf{z}^j = \mathbf{c}[\zeta_j] - \mathbf{c}[\mathscr{J}]^T A[\ ,\mathscr{J}]^{-1} A[\ ,\zeta_j] \qquad \text{for all} \quad j \in \overline{1, n-m}. \tag{45}$$

The $\mathscr{Z}$-simplex algorithm for solving (P) (assumed feasible) is now described.

$\mathscr{Z}$-SIMPLEX ALGORITHM

input: $A \in R_m^{m \times n}$, $\mathbf{b} \in R^m$, $\mathbf{c} \in R^n$;
$\mathscr{J}$ a basic subset of $\overline{1, n}$;
$\mathbf{x}$ the $\mathscr{J}$-b.f.s. of (P);
$\mathscr{Z} = \{\mathbf{z}^1, \mathbf{z}^2, \ldots, \mathbf{z}^{n-m}\}$ the $\mathscr{J}$-canonical basis of $N(A)$, with the list $\zeta = \{\zeta_1, \zeta_2, \ldots, \zeta_{n-m}\}$.

output: "current $\mathbf{x}$ is optimal" or "(P) is unbounded".

begin

ITERATION:

if $\mathbf{c}^T \mathbf{z}^j \leqslant 0$ for all $j \in \overline{1, n-m}$ **then** "current $\mathbf{x}$ is optimal"; **stop**.

else select $p \in \overline{1, n-m}$ with $\mathbf{c}^T \mathbf{z}^p > 0$;

if $\mathbf{z}^p[i] \geqslant 0$ for all $i \in \mathscr{J}$ **then** "(P) is unbounded"; **stop**.

else $\theta_p := \min\left\{ -\dfrac{\mathbf{x}[i]}{\mathbf{z}^p[i]} : i \in \mathscr{J}, \mathbf{z}^p[i] < 0 \right\}$;

(update $\mathscr{J}$) **select** q from among

$$q \in \arg\min\left\{ -\frac{\mathbf{x}[i]}{\mathbf{z}^p[i]} : i \in \mathscr{J}, \mathbf{z}^p[i] < 0 \right\};$$

$$\mathscr{J} := (\mathscr{J} \setminus \{q\}) \cup \{\zeta_p\};$$

(update $\mathbf{x}$) $\mathbf{x} := \mathbf{x} + \theta_p \mathbf{z}^p$;

(update $\mathscr{Z}$) **procedure** ADJACENT $N(A)$ BASIS;

go to ITERATION

end

As in the simplex algorithm, there are here two selections, of ζ_p ("entering") and q ("leaving"). These selections should use an *anticycling rule* (e.g.

[8], [9] and [5]). Except for that, the $\mathscr{X}$-simplex algorithm is a faithful implementation of Theorem 4.

The *convex simplex algorithm* [24] can be similarly described. Here the problem is

$$\max\{f(\mathbf{x}) : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0}\}$$

with a convex function f, so that if an optimal solution exists, then one of the b.f.s.'s is optimal. The required change in the $\mathscr{X}$-simplex algorithm is replacing $\mathbf{c}$, at each iteration, by $\nabla f(\mathbf{x})$, the gradient of f at the current solution $\mathbf{x}$.

The dual method [for solving $(\tilde{\mathrm{D}})$, assumed feasible] is now described analogously in terms of canonical bases $\mathscr{Y}$ of $R(A^T)$.

$\mathscr{Y}$-DUAL METHOD

input: $A \in R_m^{m \times n}$, $\mathbf{c} \in R^n$;
$\tilde{\mathbf{x}}$ any solution of $A\mathbf{x} = \mathbf{b}$ (**comment:** $\tilde{\mathbf{x}}$ is in lieu of $\mathbf{b}$);
$\mathscr{J}$ a basic subset of $\overline{1,n}$;
$\mathbf{y}$ the $\mathscr{J}$-b.f.s. of $(\tilde{\mathrm{D}})$;
$\mathscr{Y} = \{\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^m\}$ the $\mathscr{J}$-canonical basis of $R(A^T)$, with the list $v = \{v_1, v_2, \ldots, v_m\}$.

output: "current $\mathbf{y}$ is optimal" or "$(\tilde{\mathrm{D}})$ is unbounded".

begin

ITERATION:

if $\tilde{\mathbf{x}}^T\mathbf{y}^j \geqslant 0$ for all $j \in \overline{1,m}$ **then** "current $\mathbf{y}$ is optimal"; **stop.**

else select $p \in \overline{1,m}$ with $\tilde{\mathbf{x}}^T\mathbf{y}^p < 0$;

if $\mathbf{y}^p[i] \geqslant 0$ for all $i \in \mathscr{J}^c$ **then** "$(\tilde{\mathrm{D}})$ is unbounded"; **stop.**

else $\theta_p := \min\left\{ -\dfrac{\mathbf{y}[i] - \mathbf{c}[i]}{\mathbf{y}^p[i]} : i \in \mathscr{J}^c, \mathbf{y}^p[i] < 0\right\}$;

(update $\mathscr{J}$) **select** q from among

$$q \in \arg\min\left\{ -\frac{\mathbf{y}[i] - \mathbf{c}[i]}{\mathbf{y}^p[i]} : i \in \mathscr{J}^c, \mathbf{y}^p[i] < 0\right\};$$

$\bar{\mathscr{J}} := (\mathscr{J} \setminus \{v_p\}) \cup \{q\}$;

(update $\mathbf{y}$) $\mathbf{y} := \mathbf{y} + \theta_p \mathbf{y}^p$;

(update $\mathscr{Y}$) **procedure** ADJACENT $R(A^T)$ BASIS;

go to ITERATION

end

At each iteration, both algorithms employ a pair $\{\mathbf{x}, \mathbf{y}\}$ of $\mathscr{J}$-basic solutions of $\{(\mathrm{P}), (\tilde{\mathrm{D}})\}$ respectively. One of these vectors is explicitly updated (x in the $\mathscr{X}$-simplex algorithm, y in the $\mathscr{Y}$-dual method) and is kept feasible throughout the iterations. The remaining vector is not explicitly updated, but its feasibility is checked at each iteration by the optimality criterion. These pairs satisfy, by (23) and (34), *complementary slackness*:

$$(\mathbf{y}-\mathbf{c})^T\mathbf{x}=0. \tag{46}$$

By the duality theorem of LP, $\{\mathbf{x},\mathbf{y}\}$ are optimal solutions of $\{(\mathrm{P}),(\tilde{\mathrm{D}})\}$ if both x, y are feasible solutions for their respective problems.

In the $\mathscr{X}$-simplex algorithm, the sufficient condition for optimality

$$\mathbf{c}^T\mathbf{z}^j\leqslant 0 \qquad \text{for all} \quad j\in\overline{1,n-m} \tag{47}$$

is, by (42), a check of the $(\tilde{\mathrm{D}})$-feasibility of y. If (47) holds, then y is $(\tilde{\mathrm{D}})$-feasible, and therefore $\{\mathbf{x},\mathbf{y}\}$ are optimal. If (47) does not hold, then $\{\mathbf{x},\mathscr{J},\mathscr{X}\}$ are updated. The vector y is not explicitly needed.

Analogously, in the $\mathscr{Y}$-dual method y is $(\tilde{\mathrm{D}})$-feasible throughout, and the (P)-feasibility of x is checked by the optimality condition

$$\tilde{\mathbf{x}}^T\mathbf{y}^j\geqslant 0 \qquad \text{for all} \quad j\in\overline{1,m} \tag{48}$$

[see Lemma 2, (43)]. If (48) does not hold, then $\{\mathbf{y},\mathscr{J},\mathrm{Y}\}$ are updated, but not x which is implicit. Since the method enforces *dual feasibility* and *complementary slackness*, and works towards *primal feasibility*, it is called "dual."

Both the $\mathscr{X}$-simplex and $\mathscr{Y}$-dual algorithms use compact tableaus of approximate size $m\times(n-m)$ [since only the $m\times(n-m)$ matrix X in (8) needs updating, together with the appropriate list]. Simplex algorithms with compact tableaus were developed by Chvátal [10] (where the compact tableaus are called *dictionaries*), Rockafellar [23] (*Tucker representations*), and others. Here, as in other compact algorithms, each iteration requires about $2m(n-m)$ operations (multiplications and additions). This is the number of operations in adjacent updates of canonical bases by the procedures of Section 3. The number of the remaining operations per iteration is linear in m and $n-m$. In particular, the update of the reduced costs requires about $n-m$ or m operations [the direct computation of reduced costs using the inner products (42) or (43) requires about $m(n-m)$ operations]. For example, in the $\mathscr{X}$-simplex algorithm, the reduced costs $\mathbf{c}^T\mathbf{z}^j$ are updated by the formula

$$\mathbf{c}^T\mathbf{z}^p:=\frac{\mathbf{c}^T\mathbf{z}^p}{\mathbf{z}^p[q]}, \tag{49}$$

and then, for $p \neq j \in \overline{1, n-m}$,

$$\mathbf{c}^T\mathbf{z}^j := \mathbf{c}^T\mathbf{z}^j - \mathbf{z}^j[q](\mathbf{c}^T\mathbf{z}^p). \tag{50}$$

These formulas follow from the updates of the $\{\mathbf{z}^j : j \in \overline{1, n-m}\}$ in procedure ADJACENT $N(A)$ BASIS. The updates (49), (50) require about $n - m$ operations. Analogously, about m operations are required to update the reduced costs $\tilde{\mathbf{x}}^T\mathbf{y}^j$ in the $\mathscr{Y}$-dual method.

The observation that not all vectors in a canonical basis must be updated at each iteration has been used in [2] to reduce the work in the $\mathscr{X}$-simplex algorithm.

Both paper and author benefited from constructive suggestions and criticism by Sjur Flåm, Sverre Storøy, and the referee.

REFERENCES

1 M. S. Bazaraa and J. J. Jarvis, *Linear Programming and Network Flows*, Wiley, New York, 1977.

2 A. Ben-Israel, A $\mathscr{X}$-simplex algorithm with partial updates, *BIT* 27:50–61 (1987).

3 A. Ben-Israel, Recursive Linear Programming, Report CMI-852555-6, The Chr. Michelsen Inst., N-5036 Fantoft, Norway, Aug. 1985.

4 A. Ben-Israel and S. D. Flåm, Canonical Bases and Sensitivity Analysis in Linear Programming, The Chr. Michelsen Inst., N-5036 Fantoft, Norway, Aug. 1985.

5 A. Ben-Israel and S. Storøy, On the Computation of Basic Feasible Solutions in Linear Programming, The Chr. Michelsen Inst., N-5036 Fantoft, Norway, Aug. 1985.

6 M. W. Berry, M. T. Heath, I. Kaneko, M. Lawo, R. J. Plemmons, and R. C. Ward, An algorithm to compute a sparse basis of the null space, *Numer. Math.* 47:483–504 (1985).

7 J. Bisschop and A. Meeraus, Matrix augmentation and partitioning in the updating of the basis inverse, *Math. Programming* 13:241–254 (1977).

8 R. G. Bland, New finite pivoting rules for the simplex method, *Math. Oper. Res.* 2:103–107 (1977).

9 A. Charnes, Optimality and degeneracy in linear programming, *Econometrica* 20:160–170 (1952).

10 V. Chvátal, *Linear Programming*, Freeman, New York, 1983.

11 T. F. Coleman and A. Pothen, The null space problem I. Complexity, *SIAM J. Algebraic Discrete Methods* 7:527–537 (1986).

12 G. B. Dantzig, Maximization of a linear function of variables subject to linear inequalities, in *Activity Analysis of Production and Allocation*, Cowles Commission Monograph 13 (T. C. Koopmans, Ed.), Wiley, New York, 1951, Chapter XXI.

13 G. B. Dantzig, *Linear Programming and Extensions*, Princeton U.P., Princeton, 1963.
14 P. E. Gill, W. Murray, M. A. Saunders, and M. H. Wright, Sparse matrix methods in optimization, *SIAM J. Sci. Statist. Comput.* 5:562–589 (1984).
15 D. Goldfarb and J. K. Reid, A practicable steepest-edge simplex algorithm, *Math. Programming* 12:361–371 (1977).
16 M. Kallio and E. L. Porteus, A class of methods for linear programming, *Math. Programming* 14:161–169 (1978).
17 C. E. Lemke, The dual method of solving the linear programming problem, *Naval Res. Logist. Quart.* 1:36–47 (1954).
18 B. A. Murtagh and M. A. Saunders, Large scale linearly constrained optimization, *Math. Programming* 14:41–72 (1978).
19 B. Noble and J. W. Daniel, *Applied Linear Algebra* (2nd ed.), Prentice-Hall, Englewood Cliffs, N.J., 1977.
20 A. Orden, Stationary points of quadratic functions under linear constraints, *Comput. J.* 7:238–242 (1964).
21 C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, N.J., 1982.
22 R. T. Rockafellar, The elementary vectors of subspaces in R^n, in *Combinatorial Mathematics and its Applications* (R. C. Bose and T. A. Dowling, Eds.), Univ. of North Carolina Press, Chapel Hill, 1969, pp. 104–127.
23 R. T. Rockafellar, *Network Flows and Monotropic Optimization*, Wiley, New York, 1984.
24 W. I. Zangwill, *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
25 G. Zoutendijk, *Methods of Feasible Directions*, Elsevier, Amsterdam, 1960.

Received 12 August 1985; final manuscript accepted 11 June 1987